ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia - Social and Behavioral Sciences 54 (2012) 75 – 83

Procedia Social and Behavioral Sciences

EWGT 2012

15th meeting of the EURO Working Group on Transportation

Modelling travel time reliability with the Burr distribution

Michael A P Taylor*, Susilawati

Barbara Hardy Institute and School of Natural and Built Environments, University of South Australia, Adelaide SA 5000, Australia

Abstract

This paper suggests the Burr distribution as a useful statistical model to represent travel time reliability through studies of the day-to-day variability in travel times on journeys in urban areas. This distribution has a flexible shape and the ability to describe the very long upper tails (and hence significant skewness) seen in observed distributions of travel time variations. This result provides a means to study travel time reliability in more detail, and in line with recent research on indices of travel time reliability. The Burr distribution is algebraically tractable, which means that percentile values can be computed directly. In this way various travel time reliability metrics, such as the FHWA Buffer Index and the Delft skewness parameters can be computed from the fitted Burr parameters. This opens the way to the inclusion of reliability as a consideration in economic analysis, using a metric that can better reflect the nature of travel time variability. A case study example, using real world data, illustrates the proposed computational method for the travel time reliability metrics.

Keywords: Travel time; travel time reliability; reliability metrics; Burr distribution; maximum likelihood estimation; travel time database

1. Introduction

Travel time reliability is now a key indicator for traffic performance assessment. There is an increasing demand to include reliability in the evaluation and appraisal of transport projects and programs, as a separate factor in its own right. A major issue in this process is the selection of appropriate statistical distributions and metrics to model and represent travel time reliability. The initial choices were the log-normal distribution and the standard deviation of observed individual travel times. The problem with these choices is the strong skew to the

* Corresponding author. Tel.: +61-88-302-1861; fax: +61-88-302-1880.
E-mail address: map.taylor@unisa.edu.au.

 doi:10.1016/j.sbspro.2012.09.727

upper tail often observed in empirical travel time distributions. While the log-normal distribution has a skew to the upper tail, this skew is limited by the mathematical form of its probability density function to the extent that it cannot be 'bent enough' to capture the spread of observations seen in the upper tail. The standard deviation is a symmetric measure, and a directional measure of dispersion may be more appropriate. The question is then to find a suitable distribution and metrics for this purpose.

As described in [1], previous studies have suggested that the travel time variability distribution might follow either normal or lognormal forms. A comprehensive data analysis involving the assessment of continuous and longitudinal travel time data for a number of routes and corridors in metropolitan Adelaide, Australia demonstrates that the observed distributions appear to be more complex than previously assumed. Not only is there is evidence of very long upper tails in the observed data sets, there is even bimodality in some cases. This research examines the use of alternative statistical distributions for longitudinal variations of travel times. Of these alternatives, the Burr distribution emerges as a useful model representing link and route level travel time variability distributions. This distribution also offers computational advantages for the use of metrics (e.g. the Delft *skew-width* indices, Van Lint and Van Zuylen [2]) based on percentiles.

2. Travel time reliability metrics

The standard deviation and the coefficient of variation have been the common parameters adopted to describe how travel times vary [3]. There is, however, some discussion as the appropriateness of the standard deviation, given the common observation that the distributions of variability in travel times are significantly skewed to the upper tail [1, 4]. Some studies (e.g. [5] argue that variability may be better represented by considering a percentile value, such as the 95th percentile travel time, as this reflects a logical measure of 'being late' (the 95th percentile implies a probability of one in 20 of exceeding that travel time value, which for a commuter might mean being late for work once a month).

Some alternative travel time reliability metrics have thus been proposed. FHWA [5] introduced the buffer time (BT_t) to represent the additional time above the average travel time ($\bar{t}$) required for on-time arrival. The buffer time is the difference between the 95th percentile travel time (t_{95}) and the mean travel time:

$$BT_t = t_{95} - \bar{t}$$

The Buffer Time Index (BTI_t) is then defined as

$$BTI_t = \frac{BT_t - \bar{t}}{\bar{t}} = \frac{t_{95}}{\bar{t}} - 1 \tag{1}$$

Additionally, Van Lint and Van Zuylen [2] proposed the so-called *skew-width methods*. The skewness of the travel time λ^{skew} is defined as the ratio of difference of 90th percentile and 50th percentile travel time and the differences between 50th percentile and 10th percentile travel time. This ratio relates the range of the 40 per cent of observations above the median to the range of the 80 per cent below it, and thus indicates the level and direction of skewness in the data. The width of travel time λ^{var} is defined as the ratio of the differences of 90th percentile and 50th percentile travel time and the 50th percentile travel time. The equation of the skewness and the width travel time metrics are as follows:

$$\lambda^{skew} = \frac{t_{90} - t_{50}}{t_{50} - t_{10}} \tag{2}$$

$$\lambda^{\text{var}} = \frac{t_{90} - t_{10}}{t_{50}} \tag{3}$$

As an example of the use of the standard deviation (s) of individual travel times as a reliability metric, Eliasson (2006) developed a model for estimating s in terms of mean travel time, link length (L) and free flow travel time. This model is

$$s = \rho \lambda_{TOD} \lambda_{SPD} L^{\kappa} \bar{t}^{\gamma} \left(\frac{\bar{t}}{t_f} - 1 \right)^{\omega}$$

where λ_{TOD} and λ_{SPD} are dummy variables representing time of day and the speed limit, ρ is a constant, and κ, γ and ω are estimated parameters.

Herman and Lam [7] and Richardson and Taylor [8] noted the skewness to the right in their observed distributions of travel time variations, and suggested that, on the basis of the data sets available to them, the log-normal distribution could provide a useful model for the distributions. More recent work using substantial new data (the Adelaide Longitudinal Travel Time Reliability Database, ALTTRDAT) has indicated that the log-normal distribution still does not have the ability to fully represent the long upper tails in the observed data [1]. Using research from recent reliability engineering studies, we noted the advantages of the Burr Type XII distribution (subsequently termed the Burr distribution) in modeling distribution with long upper tails. Susilawati et al [1] then showed that the Burr distribution is one continuous distribution that fits much of the observed travel time variability data.

The Burr distribution is also well known in actuarial theory, where it has found a place in modeling distributions of insurance claims. It was developed by Burr [9] for the express purpose of fitting a cumulative distribution function (cdf) to a diversity of frequency data forms. In its basic form it has two parameters, $c > 0$ and $k > 0$, which are shape parameters.

3. Burr distribution

The probability density function (pdf) $f(x, c, k)$ of the (2-parameter) Burr distribution is

$$f(x,c,k) = ckx^{c-1}(1+x^c)^{-(k+1)}$$

where $x > 0$, $c > 0$ and $k > 0$. The cdf $F(x, c, k)$ is given by

$$F(x,c,k) = 1-(1+x^c)^{-k} \tag{4}$$

The distribution has some interesting statistical properties [10]. In the first instance the rth moment of the distribution $E(x^r)$ will only exist if $ck > r$, in which case

$$E(x^r) = \mu'_r = \frac{k\Gamma(k-\frac{r}{c})\Gamma(\frac{r}{c}+1)}{\Gamma(k+1)}$$

where $\Gamma(y)$ is the mathematical Gamma function. The mean of the Burr distribution is thus

$$E(x) = \bar{x} = \frac{k\Gamma(k - \frac{1}{c})\Gamma(\frac{1}{c} + 1)}{\Gamma(k+1)} \tag{5}$$

In addition, the modal value of x is x_m, which is given by

$$x_m = \left[\frac{c-1}{ck+1}\right]^{1/c}$$

but x_m will only exist if $c > 1$. [If $c \leq 1$, then the distribution is L-shaped.] The Burr distribution therefore has a flexible shape and is well behaved algebraically.

4. Derivation of metrics

A further advantage of the Burr distribution is its algebraic tractability, which means (for instance) that percentile values can be computed directly. In this way various travel time reliability metrics, such as the Buffer Time Index and the Delft skewness parameters can be computed in terms of the fitted Burr parameters. As described above these metrics are based on percentile values for the travel time distribution. This opens the way to the inclusion of reliability as a consideration in economic analysis, using a metric that can better reflect the nature of travel time variability (e.g. the long upper tail of the distribution).

Percentiles for the Burr distribution may be computed using the following approach. Given the cdf defined by equation (4), we can solve for a given value of *F(x, c, k)* for percentile *P*, i.e.

$$P = 1 - (1 + x^c)^{-k}$$

from which

$$x_P = \sqrt[c]{(1-P)^{-1/k} - 1} \tag{6}$$

The median is then

$$x_{50} = \sqrt[c]{2^{1/k} - 1} \tag{7}$$

and the 95th percentile is

$$x_{95} = \sqrt[c]{20^{1/k} - 1} \tag{8}$$

It is also possible to define other travel time variability metrics in terms of the Burr distribution parameters (c and k). Consider the *Buffer Time Index* defined by equation (1). In terms of the scaled travel time variable $x = \alpha t$ where $\alpha > 0$ is a constant, we can write a *Buffer Index* BI_x as

$$BI_x = \frac{x_{95}}{\bar{x}} - 1$$

and we already know x_{95} and $\bar{x}$ as functions of c and k – see equations (5) and (8), so that

$$BI_x = \frac{\sqrt[c]{20^k - 1}}{\dfrac{k\Gamma(k - \frac{1}{c})\Gamma(\frac{1}{c} + 1)}{\Gamma(k+1)}} - 1 \qquad (9)$$

which is a relatively simple function of c and k. Using similar reasoning, we can define the Delft travel time reliability metrics 'skewness of travel time' (λ^{skew}) and 'width of travel time' (λ^{var}) in terms of the Burr parameters, given that these two metrics are both defined in terms of percentiles. The metrics are defined in equations (2) and (3). Given x_{50} as defined by equation (7) and that $x_{90} = \sqrt[c]{10^{1/k} - 1}$ and $x_{10} = \sqrt[c]{\left(\frac{10}{9}\right)^{1/k} - 1}$ it follows that the Delft *skewness metric* is

$$\lambda^{skew} = \frac{\sqrt[c]{10^{1/k} - 1} - \sqrt[c]{2^{1/k} - 1}}{\sqrt[c]{10^{1/k} - 1} - \sqrt[c]{\left(\frac{10}{9}\right)^{1/k} - 1}} \qquad (10)$$

and the Delft *width metric* is

$$\lambda^{\mathrm{var}} = \frac{\sqrt[c]{10^{1/k} - 1} - \sqrt[c]{\left(\frac{10}{9}\right)^{1/k} - 1}}{\sqrt[c]{2^{1/k} - 1}} \qquad (11)$$

This latter metric is the ratio of the middle 80 per cent of the data to the median travel time. It has proven useful for describing travel time reliability on urban arterial roads, and as described in Susilawati [11] can be used to related travel time reliability to traffic engineering performance measures such as the degree of saturation.

5. Application

ALTTRDAT contains day-by-day variations in travel time on several routes through the Adelaide metropolitan area. Collection of the travel time data commenced in 2007 on the first route, the Glen Osmond Road route, which is shown in Figure 1. Other routes have been progressively added since then and there are now six separate routes. Data is collected using second-by-second GPS data, providing a rich source of travel time data for each route and for defined sections (links) within it. Observations over many days, weeks and months enable the development of longitudinal travel time distributions reflecting the day-by-day variations in travel times experienced by commuters, and thus their experiences and perceptions about a given route. The Glen Osmond Road route is shown in red in Figure 1. It is divided into seven links, and Links 3 and 5 are shown on the map.

The Glen Osmond Road route is 5.920 km long. The first 2.204 km (from the start, at the southern end, to Fullarton Road – see Figure 1) is a four-lane single-carriageway arterial road passing through a predominantly (low density) residential area. The section from Fullarton Road to Greenhill Road (Link 3 in Figure 1) is also four-lane single-carriageway arterial road, but through a mixed commercial/retail area. This link is 1.064 km long. Greenhill Road to South Terrace (0.776 km) is a four-lane divided arterial road through parklands. A 60 km/h speed limit applies on all of these preceding route sections, and clearway restrictions apply in the morning peak (the time of day for the travel time runs) so kerbside parking is prohibited at that time. The remainder of the route lies in the Adelaide central business district (CBD) and a 50 km/h speed limit applies there. Link 5 (see Figure 1) is along Hutt Street, a popular cafe/restaurant street, which is four lane divided carriageway with angle parking on each side. This route section is 0.791 km long. The remaining 1.085 km of the route continues through the CBD on either 2 lane single carriageway road or four-lane single carriageway road (both with separate parking lanes). There are 17 sets of traffic signals along the route, and all links except the final one end at a traffic signal. The final link ends at a carpark entrance.

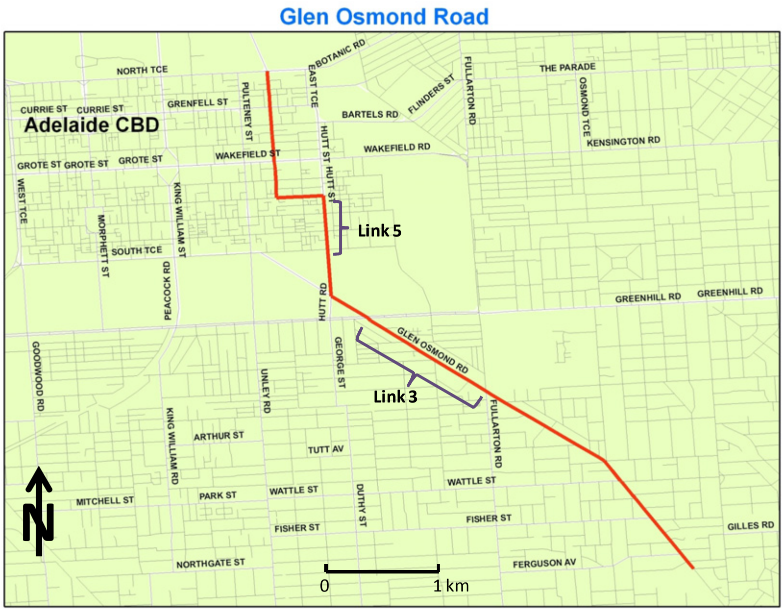

Fig. 1. Map showing the Glen Osmond Road travel time route in the Adelaide Longitudinal Travel Time Reliability Database (ALTTRDAT)

The GPS data for each travel time run form a data stream including second-by-second location, time and speed. The data may be used to construct speed-time (v-t) and speed-distance (v-x) profiles and a trajectory (x-t) diagram for the run. This is a rich source of data defining the progression of the test vehicle along the route. Figure 2 (a), (b) and (c) show v-t, v-x and x-t plots for a sample run. Section (link) travel times may then be extracted from these result – and the definition of each section is flexible in that the data can be analysed for different section configurations as required.

Table 1 shows example summary results from one year of travel time data for the Glen Osmond Road route, including the total route travel time and data for Links 3 and 5 (see Figure 1). The table includes mean, standard deviation and coefficient of variation of the section travel times, as well as percentile values (t10, t50, t90 and t95).

Table 1. Maximum likelihood estimates and goodness of fit results for the Burr distribution fitted to Glen Osmond Road travel time distributions

Data set	Link results					Percentiles (min)			
	Length (km)	Mean tt (min)	St dev (min)	Coeff of variation	Mean speed (km/h)	t10	t50	t90	t95
Whole route	5.920	14.31	2.78	19.4%	24.8	10.93	13.53	18.44	20.90
Link 3	1.064	2.45	1.17	46.4%	26.0	1.88	2.19	3.27	3.89
Link 5	0.791	3.40	1.66	48.7%	14.0	1.95	3.03	5.13	6.04

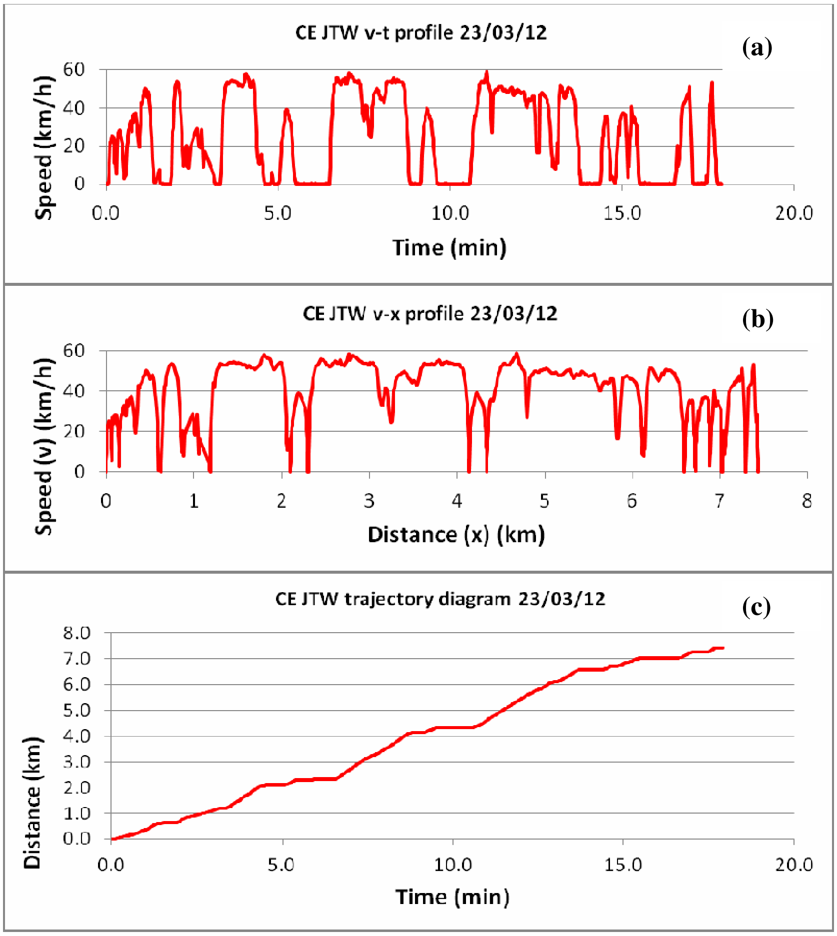

Fig. 2. Speed-time (a), speed-distance (b) and trajectory (c) diagrams for a GPS travel time run along the Glen Osmond Road travel time route

The Burr distribution has been fitted to the data for the complete route, and to individual links in the route, using maximum likelihood to estimate the distribution parameters. Table 2 shows the estimated parameters and the goodness of fit test results for the route and for Links 3 and 5. Some travel time metrics – Buffer Time Index and the Delft metrics – calculated using the estimated Burr parameters are then shown Table 3.

Values of λ^{skew} exceeding one indicate a skew to the right (upper tail) of the distribution, and the higher the value, the greater the skew. λ^{var} Indicates the spread of the data about the median, and a value of 0.6908 indicates that the magnitude of the range of the middle 80 per cent of the data is about 69 per cent of the median value As a typical example of the fitted distributions, Figure 3 shows the observed histogram and Burr pdf for Link 5. The characteristic long tail in the observed data, reflecting the inherent skewness in the data, is clearly seen.

Table 2. Maximum likelihood estimates and goodness of fit results for the Burr distribution fitted to Glen Osmond Road travel time distributions

	Maximum likelihood estimates		Komolgorov-Smirnov test of goodness of fit			
Data set	c	k	D-stat	0.05 critical value	p-value	Result
Whole route	12.40	0.446	0.0540	0.1018	0.6558	Accepted
Link 3	44.71	0.090	0.0578	0.1050	0.5734	Accepted
Link 5	5.344	0.671	0.0379	0.1018	0.9512	Accepted

Table 3. Computed travel time reliability metrics using Burr distribution maximum likelihood estimates for Glen Osmond Road travel time distributions

	Fitted Burr parameters			Travel time reliability metrics		
Data set	c	k	ck	FHWA Buffer Index	Delft skewness (λ^{skew})	Delft width (λ^{var})
Whole route	12.40	0.446	5.5308	0.4609	1.8934	0.5547
Link 3	44.71	0.090	4.0239	0.5858	3.4381	0.6348
Link 5	5.344	0.671	3.5844	0.8363	1.9314	0.6908

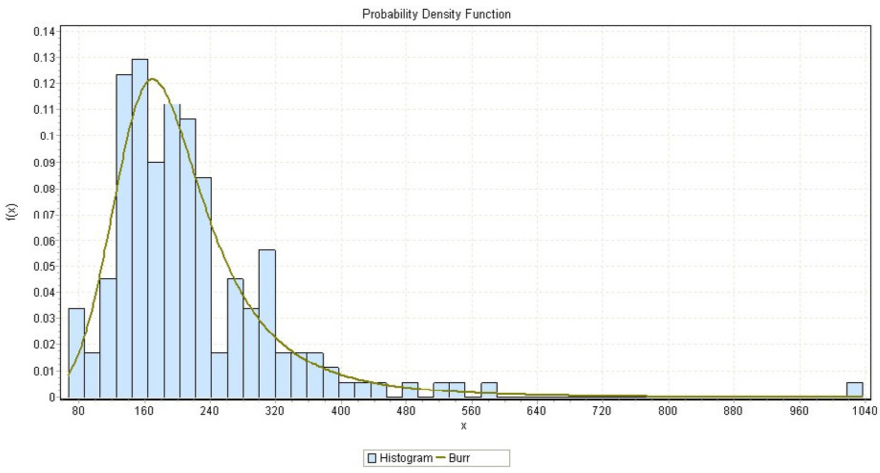

Fig. 3. Observed histogram and fitted Burr distribution for Link 5 in the Glen Osmond Road travel time route (Adelaide Longitudinal Travel Time Reliability Database)

6. Conclusions

This paper has suggested a new statistical distribution, the Burr distribution, as a useful statistical model to represent travel time reliability through studies of the day-to-day variability in travel times on journeys in urban areas. This distribution has a flexible shape and the ability to describe the very long upper tails (and hence significant skewness) seen in observed distributions of travel time variations. This is a useful result because it provides a means to study travel time reliability in more detail, and in line with recent research on suitable indices to measure reliability.

A further advantage of the Burr distribution is its algebraic tractability, which means (for instance) that percentile values can be computed directly. In this way various travel time reliability metrics, such as the FHWA Buffer Index and the Delft skewness parameters can be computed directly in terms of the fitted Burr parameters. These metrics are all based on percentile values for the travel time distribution. This opens the way to the inclusion of reliability as a consideration in economic analysis, using a metric that can better reflect the nature of travel time variability. A case study example, using real world data, was used to illustrate the proposed computational method for the travel time reliability metrics.

7. References

1. Susilawati, Taylor, M.A.P. & Sekhar, S.V.C. (2012). Distributions of travel time variability on urban roads. *Journal of Advanced Transportation* (in press, DOI: 10.1002/atr.192).
2. Van Lint, J.W.C. and Van Zuylen, H.J. (2005). Monitoring and predicting travel time reliability: using width and skew of day-to-day travel time distribution. *Transportation Research Record*, 1917, 54-62.
3. Bates, J., Polak, J., Jones, P. & Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research E,* 37, 191-229.
4. Fosgereau, M. & Karlstrom, A. (2010) The value of reliability. *Transportation Research B*, 44 (1), 38-49.
5. FHWA. (2006) Travel time reliability: making it there on time, all the time. Federal Highway Administration US DOT, Washington DC. http://ops.fhwa.dot.gov/publications/tt_reliability/TTR_Report.htm
6. Eliasson, J. (2006). The relationship between travel time variability and road congestion. Working Paper, Centre for Transport Studies, Royal Institute of Technology, Stockholm.
7. Herman, R. & Lam, T. (1974). Trip time characteristics of journeys to and from work. In D.J.Buckley (Ed.) *Transportation and Traffic Theory* (pp.57-85). Sydney: A.H. and A.W. Reed.
8. Richardson, A.J. & Taylor, M.A.P. (1978) Travel time variability on commuter journeys. *High Speed Ground Transportation Journal*, 6, 77-99.
9. Burr, I.W. (1942). Cumulative frequency functions. *The Annals of Mathematical Statistics*, 13, 215-232.
10. Tadikamalla, P.R. (1980). A look at the Burr and related distributions. *International Statistical Review*, 48, 337-344.
11. Susilawati (2011). Modelling urban arterial road travel time variability. Unpublished PhD thesis, University of South Australia, Adelaide